

Working Paper 93-45
Statistics and Econometrics Series 27
December 1993

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

FORECASTING GROWTH WITH TIME SERIES MODELS

Daniel Peña*

Abstract

This paper compares the structure of three models for estimating future growth in a time series. It is shown that a regression model gives minimum weight to the last observed growth and maximum weight to the observed growth in the middle of the sample period. A first order integrated ARIMA model, or $I(1)$ model, gives uniform weights to all observed growths. Finally, a second order integrated ARIMA model gives maximum weights to the last observed growth and minimum weights to the observed growths at the beginning of the sample period.

Key Words

ARIMA Models; Integrated Processes; Regression; Stationary Processes.

*Departamento de Estadística y Econometría, Universidad Carlos III de Madrid.

FORECASTING GROWTH WITH TIME SERIES MODELS

by

Daniel Peña

Department of Statistics and Econometrics

Universidad Carlos III de Madrid

SUMMARY

This paper compares the structure of three models for estimating future growth in a time series. It is shown that a regression model gives minimum weight to the last observed growth and maximum weight to the observed growth in the middle of the sample period. A first order integrated ARIMA model, or $I(1)$ model, gives uniform weights to all observed growths. Finally, a second order integrated ARIMA model gives maximum weights to the last observed growth and minimum weights to the observed growths at the beginning of the sample period.

Key words: ARIMA models; Integrated processes; Regression; Stationary processes.

1. INTRODUCTION

An important problem in modeling economic time series is forecasting the future growth of a given time series. Assuming that a linear model is appropriate for the data, the procedures most often used are as follows: (i) detrend the observed data by regressing the observations on time, and use the residuals from this regression to build a stationary time series model. The series is forecasted by adding the values of the deterministic future trend and the forecast of the stationary residual; (ii) differentiate the series, test for unit roots and if the series is assumed to be integrated of order one (I(1)) build a stationary ARMA model in the first difference of the series. Typically models built in this way include a constant for many economic time series; (iii) differentiate twice the series and build the ARMA model on the second difference of the process that it is assumed to be I(2), then, in most cases the I(2) model does not include a constant term. The decision between these three procedures should be done by testing the number of unit roots in the time series model. However, the available tests are not very powerful, specially for short time series, and, therefore, it is important to understand the consequences of using these models.

Let Z_t be the time series data and let us call $b_t = Z_t - Z_{t-1}$ the observed growth at time t , it is shown in this paper that the estimate of future growth by the three procedures can be written as

$$\hat{\beta}_f = \sum \omega_t b_t$$

where the coefficients ω_t are a weighting function, that is $\omega_t > 0$ and $\sum \omega_j = 1$. Section 2 of this paper proves that linear regression gives minimum weights to the last observed growth and maximum weights to the observed growth in the middle of the sample. Section 3 shows that an I(1) model with a constant term gives a uniform weight throughout the sample, that is $\omega_t = n^{-1}$. Finally, Section 4 shows that an I(2) model gives maximum weight to the last observed growth and minimum to the oldest values. As this last procedure seems intuitively more appropriate for economic time series, it is argued in Section 5 that it should receive a careful consideration for forecasting growth in economic time series.

2. REGRESSION ON TIME

Let us call Z_t the observed time series and let us assume for the sake of simplicity that the sample size is $n=2m+1$. Let $t=\{-m, \dots, 0, \dots, +m\}$. Then the least squares estimator of the slope in the regression on time

$$Z_t = \beta_0 + \beta_1 t + u \quad (2.1)$$

is given by

$$\hat{\beta}_1 = \frac{\sum t Z_t}{\sum t^2} = \left[2 \sum_{i=1}^m i^2 \right]^{-1} \sum_{i=1}^m i (Z_i - Z_{-i}). \quad (2.2)$$

Calling $b_t = Z_t - Z_{t-1}$, for $t = -m+1, \dots, m$, the observed growth at each period, we note that

$$Z_t - Z_{-t} = \sum_{j=-t+1}^{j=t} b_j$$

and, after some straightforward manipulations that are shown in the appendix, the estimate of the slope can be written as

$$\hat{\beta}_1 = \sum_{j=1}^m \omega_j (b_j + b_{1-j}) \quad (2.3)$$

where the weights ω_j are given by

$$\omega_j = \frac{3(jm)(m-j+1)}{2(2m+1)m(m+1)}, \quad j = 1, \dots, m \quad (2.4)$$

and add up to one. Therefore the estimated growth $\hat{\beta}_1$ is a weighted mean of all the observed growths b_j , such that the maximum weights are given to b_1 and b_0 , that correspond to the observed growth in the middle of the sample period, and the minimum weights are given to b_m and b_{1-m} , the first and last observed growth.

The estimator (2.3) has an interesting interpretation. In the assumption that the linear model (2.1) hold, the $2m$ values b_t ($t = -m+1, \dots, m$) are unbiased estimates for β . These estimates are correlated and have covariances

$$\begin{aligned} \text{Cov}(b_t, b_{t+1}) &= E[(b_t - \beta)(b_{t+1} - \beta)] = E[(u_t - u_{t-1})(u_{t+1} - u_t)] = -\sigma^2 \\ \text{Cov}(b_t, b_{t+j}) &= 0 \quad j \geq 1. \end{aligned}$$

Therefore, the covariance matrix of these $2m$ estimates is the Toeplitz matrix:

$$V = \begin{bmatrix} 2\sigma^2 & -\sigma^2 & 0 & \dots & 0 \\ -\sigma^2 & 2\sigma^2 & . & & \\ . & . & . & . & \\ . & . & . & . & -\sigma^2 \\ 0 & \dots & \dots & -\sigma^2 & 2\sigma^2 \end{bmatrix}. \quad (2.5)$$

It is easy to show (Newbold and Granger, 1974) that given a vector $\hat{\theta}$ of unbiased estimators of a parameter θ with covariance matrix V , the best (in the mean squared sense) linear unbiased estimator of θ is given by

$$\hat{\theta}_T = (1'V^{-1}1)^{-1}(1'V^{-1}\hat{\theta}) \quad (2.6)$$

where $1' = (1 \ 1 \ \dots \ 1)$. Now, the inverse of the Toeplitz matrix (2.5) has been studied by Shaman (1969) who obtained the exact inverse of a first order moving average process. As V can be interpreted as the covariance matrix of a non-invertible ($\theta=1$) first order moving average process, then $V^{-1} = \{v_{ij}\}$, will be given by

$$v_{ij} = \frac{i(2m-j+1)}{2n+1} \quad j \geq i, \quad i=1, \dots, 2m,$$

and $v_{ij} = v_{ji}$. Therefore

$$V^{-1} = \frac{1}{(2m+1)} \begin{bmatrix} 2m & 2m-1 & 2m-2 & \dots & 1 \\ 2m-1 & 2(2m-1) & 2(2m-2) & \dots & 2 \\ 2m-2 & 2(2m-2) & 3(2m-2) & \dots & 3 \\ \vdots & \vdots & \vdots & & \\ 2 & 4 & 6 & \dots & 2m-1 \\ 1 & 2 & 3 & \dots & 2m \end{bmatrix} \quad (2.7)$$

It is proved in the appendix using this result, that the estimator (2.3) can also be obtained results by applying (2.6) to the unbiased but correlated estimates b_i .

Suppose now that an ARMA model is fitted to the residuals of the regression model (2.1). Then, the equation for the h steps ahead forecast will be

$$\hat{Z}_t(h) = \hat{\beta}_0 + \hat{\beta}_1 h + \hat{\eta}_t(h)$$

where $\hat{\eta}_t(h)$ is the forecast of the zero mean stationary process fitted to the residuals. As for a stationary process the long run forecast converges to the mean, $\hat{\eta}_t(h) \rightarrow 0$, and the parameter $\hat{\beta}_1$ will be the long-run estimated growth of the time series.

3. FORECASTING GROWTH WITH AN I(1) MODEL

The ARIMA approach in modelling time series with trend is to differentiate the data and then, fit an stationary ARMA process. Assuming that a difference is enough to obtain a stationary series, that is the series is integrated of order one or I(1), the fitted model is

$$\nabla Z_t = \beta + n_t \quad (3.1)$$

where n_t follows an ARMA model

$$n_t = \sum \psi_i a_{t-i} \quad (3.2)$$

the process $\{a_t\}$ is a gaussian white-noise process and the series $\{\psi_i\}$ converge so that n_t is a zero mean stationary process. Calling V to the covariance matrix of n_t , the estimate of β in (3.1) is given by the generalized least squares estimator

$$\hat{\beta} = (1'V^{-1}1)^{-1}(1'V^{-1}b) \quad (3.3)$$

where the vector b has components $b_t = \nabla Z_t$. Assuming that n_t is stationary and invertible it is well known (see Fuller 1976) that $\bar{b} = 1/n \sum b_t$ is asymptotically unbiased for β with variance σ^2/n .

When n is large, the expected growth h periods ahead is given by

$$\beta_T(h) = E[Z_{T+h} - Z_{T+h-1}]$$

and it will be estimated by

$$\hat{\beta}_T(h) = \hat{\beta} + \hat{n}_t(h)$$

where $\hat{n}_t(h)$ is the h -step ahead forecast of the stationary process n_t . As for h large the $\hat{n}_t(h)$ will go to zero the mean value forecast, the long-run growth will be estimated by $\hat{\beta}$. As

$$\hat{\beta} = \frac{1}{n} \sum b_t = \frac{1}{n} (Z_n - Z_1)$$

the long run growth will be estimated simply by using the first and last observed values. Also, this estimate can be interpreted as a weighted average with uniform weighing of the observed growths b_t .

4. FORECASTING GROWTH WITH AN I(2) MODEL

Some economic time series, required differencing twice to obtain a stationary model. Then, the series is called integrated of order two or I(2), and the model used is

$$\nabla^2 Z_t = n_t \quad (4.1)$$

where

$$n_t = \sum \psi_i a_{t-i} \quad (4.2)$$

and the process $\{a_t\}$ is a gaussian white-noise process. The series $\{\psi_i\}$ converge so that n_t is a zero mean stationary and invertible process. The h step-ahead forecast from model (4.1) can be written

$$\hat{Z}_t(h) = \hat{\beta}_0^{(t)} + \hat{\beta}_1^{(t)}h + \hat{n}_t(h) \quad (4.3)$$

where $\hat{\beta}_0^{(t)}$ and $\hat{\beta}_1^{(t)}$ depend on the origin of the forecast and $\hat{n}_t(h)$ is the h step-ahead forecast of the zero mean stationary process. Again, as the forecast $\hat{n}_t(h)$ will go to zero, the long-run growth will be estimated by $\hat{\beta}_1^{(t)}$. To understand the structure of $\hat{\beta}_1^{(t)}$ let us consider first the simplest case in which n_t follows an MA(1) process, $n_t = (1-\theta B)a_t$. Then the forecast for any lag h is given by

$$\hat{Z}_t(h) = \hat{\beta}_0(t) + \hat{\beta}_1^{(t)}h \quad (4.4)$$

because $\hat{n}_t(1)$ is a constant. Let us obtain the form of $\hat{\beta}_1^{(t)}$ as a function of the observed growths ∇Z_t . Assuming that the origin is $t=T-1$, then we can obtain $\hat{\beta}_0^{(T-1)}$ in (4.4) using the two forecast $\hat{Z}_{T-1}(1)$ and $\hat{Z}_{T-1}(2)$ as follows:

$$\begin{aligned}\hat{Z}_{T-1}(1) &= Z_{T-1} + \nabla Z_{T-1} - \theta a_{T-1} = \hat{\beta}_0^{(T-1)} + \hat{\beta}_1^{(T-1)} \\ \hat{Z}_{T-1}(2) &= 2\hat{Z}_{T-1}^{(1)} - Z_{T-1} = \hat{\beta}_0^{(T-1)} + 2\hat{\beta}_1^{(T-1)}\end{aligned}$$

and subtracting the first equation from the second

$$\hat{\beta}_1^{(T-1)} = \nabla Z_{T-1} - \theta a_{T-1} = b_{T-1} - \theta(1-\theta B)^{-1} \nabla b_{T-1} ;$$

which leads to

$$\hat{\beta}_1^{(T-1)} = (1-\theta) [b_{T-1} + \theta b_{T-2} + \theta^2 b_{T-3} \dots] \quad (4.5)$$

that is, the forecasted future growth is an exponentially weighted average of past observed growths.

In general it is easy to show that

$$\hat{\beta}_1^{(T-1)} = \sum a_j b_{T-j}$$

where the a_j coefficients depend on the moving average structure of the process and behave like the $\pi(B) = \psi(B)^{-1}$ structure of the process.

5. CONCLUSION

We have compared in this paper three time series models. The three models forecast future growth by using a weighted average of the observed growths in the sample. Linear regression gives minimum weight to the last observed growth and maximum weight to the center of the sample period. This implies that, for instance, if we use this method to forecast next year gross national product (GNP) with a sample of 40 data, we are saying that the most informative item to forecast 1994 growth is the growth in 1974, whereas the last observed growth in 1993 receives a weight equal to the one in 1954. If we use an I(1) model, the growth is forecasted by using a uniform weighting in all the years in the sample. In the GNP example the observed 1993 growth is as relevant as the one observed in 1960 or 1965 for

forecasting 1994 growth. The logical requirement that the most relevant year to forecast GNP growth are the last observed growth is only accomplished by using an $I(2)$ model. In particular, an ARIMA (0,2,1) model leads to an exponentially weighting of last observed growths.

Many econometric papers and some well-known time series books (see for instance Brockwell and Davis (1987), pp 25) used least squares regression on time as an alternative to differencing for removing a trend in time series. However, the logical implications of both procedures are seldom analyzed. It is important to stress that if a series follows an $I(2)$ model but we detrend it by least squares regression on time, the residuals from this fit do not provide, in general, a sound ground to fit an ARMA model, and the forecast performance of the procedure may be poor.

ACKNOWLEDGEMENT

This research has been supported by Grants PB90-0266, DGICYT, and ROB91-0244 CICYT, Spain.

APPENDIX

Using that

$$\sum_{i=1}^m i^2 = \frac{(2m+1)m(m+1)}{6} \quad (\text{A.1})$$

and

$$\begin{aligned} \sum_{i=1}^m i \sum_{j=i}^{m+1-i} b_j &= (b_1+b_0) \sum_{i=1}^m i + (b_2+b_{-1}) \sum_{i=2}^m i + \dots + (b_m+b_{1-m})m \\ &= \sum_{j=1}^m \left(\sum_{i=j}^m i \right) (b_j+b_{1-j}) \end{aligned}$$

we have that

$$\hat{\beta} = \sum_{j=1}^m \omega_j (b_j+b_{1-j}) \quad (\text{A.2})$$

where

$$\omega_j = \frac{3(j+m)(m-j+1)}{(2m+1)m(m+1)}$$

and the sum of all the weights ω_j adds up to one

$$2 \sum_{j=1}^m \omega_j = \left(\sum_{i=1}^m i^2 \right)^{-1} \sum_{j=1}^m \sum_{i=j}^m i = 1.$$

On the other hand, let $1' = (1, \dots, 1)$ be a vector of $2m$ ones. Then, using (2.7)

$$1'V^{-1} = (m, (2m-1), \frac{3}{2}(2m-2), \dots, \frac{i}{2}(2m-i+1), \dots, m)$$

and

$$1'V^{-1}1 = \sum_{j=1}^{2m} \frac{i}{2} (2m-i+1) = \frac{m(2m+1)(m+1)}{3}$$

therefore, the estimate is given by

$$\hat{\beta} = \sum_{i=1}^{2m} \frac{3i(2m-i+1)}{m(2m+1)(m+1)} b_{i-m}$$

but

$$\begin{aligned} \hat{\beta} &= \sum_{i=1}^m \frac{3i(2m-i+1)}{m(2m+1)(m+1)} b_{i-m} + \sum_{j=1}^m \frac{3(m+j)(m-j+1)}{2m(2m+1)(m+1)} b_j \\ &= \sum_{j=1}^m \frac{3(m-j+1)(m+j)}{2m(2m+1)(m+1)} b_{1-j} + \sum_{j=1}^m \frac{3(m+j)(m-j+1)}{2m(2m+1)(m+1)} b_j \\ &\quad \sum_{j=1}^m \omega_j (b_j + b_{1-j}) \end{aligned}$$

in agreement with (A.2).

REFERENCES

- Box, G.E.P., and Jenkins, G.M. (1976) *Time Series Analysis, Forecasting and Control*. Holden-Day.
- Brockwell, P.J., and Davis, R.A. (1987) *Time Series: Theory and Methods*. Springer-Verlag.
- Newbold, P., and Granger, C.W.J. (1974) Experience with Forecasting Univariate Time Series and the Combination of Forecasts. *Journal of Royal Statistical Society A*, 131-146.
- Fuller, W.A. (1976) *Introduction to Statistical Time Series*. John Wiley.
- Shaman, P. (1969) On the inverse of the covariance matrix of a first order moving average. *Biometrika*, 56, 595-600.